DOCUMENT RESUME

ED 248 727 FL 014 587

AUTHOR Palmer, Ian C.
TITLE The Ethics of Test Preparation at Intensive English Language Programs.
PUB DATE Jul 84
NOTE 12p.
PUB TYPE Reports - Descriptive (141) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *English (Second Language); *Ethics; *Intensive Language Courses; *Language Proficiency; *Language Tests; Standardized Tests; *Test Coaching; Test Wiseness

ABSTRACT

It is concluded from a review of the literature and consideration of the type and objectives of tests and programs in English as a second language that, just as limiting an instructional program to what is to be tested in standardized tests would be inappropriate, so would constructing such a program without regard for what was to be tested. It is suggested that a more realistic and effective response be taken to students' needs regarding the Test of English as a Foreign Language (TOEFL). Four test preparation methods should be addressed: (1) familiarization with the test format and procedures, (2) skill-building in regular classwork and special preparation classes, (3) practice in taking tests, and (4) training in test-taking techniques based on the test's design features as used by the test creators. Potential problems include the risk of an overly test-centered program, and potential advantages include the use of student motivation for passing the test to enhance competency and the capability of answering deep student concern about the tests. It is proposed that the responsibility of the intensive English language program is not to ignore the test because of criticism of it, but to determine how to meet the students' needs for proficiency and to display this proficiency to their best advantage. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ian C. Palmer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Ethics of Test Preparation at Intensive English Language Programs

Ian C. Palmer
Language Resource Associates

ED248727

1. Introduction

Intensive English Language Programs (IELPs) have for some time been offering to their students courses in test (notably TOEFL) preparation. Despite the increased popularity of such courses, they are still seen in certain quarters as a suspect activity. My purpose is to examine the Test of English as a Foreign Language (TOEFL), its purpose and uses, student and IELP attitudes towards it, viewpoints as to its effectiveness, and what responses to this situation are desireable on the part of the IELP.

TOEFL aims to measure English proficiency rather than academic aptitude or subject matter competence (ETS (1981), p.15). It meets the educational administrator's need to "use the most valid, reliable and economical [instrument] available" (Gue & Holdaway, p.89), and to "reduce educational wastage," becoming a "valuable contribution to social engineering" (both in Ingram, p.147). Its reliability is made more rather than less credible by the extent to which it does not discriminate among native English speakers (Alderson, p.37). TOEFL is a discrete-point test of some English language skills, an approach traced by Oller (1976, pp. 277-280) to the emphasis on segmentation and classification prevalent among linguists in the 1950's and 1960's, and to the joint influence on language teaching of contrastive analysis and audio-lingual pattern drills.

Its popularity with students and administrators alike is mainly due to the conditions of standardization and security accruing on account of its being administered by Educational Testing Service, and its incomparable "availability both to the foreign applicant and to American institutions" (Chase & Domino, p.552). It is widely thought of as the ultimate standard of evaluation of the efficiency of the IELP: "External evaluation ... can be made in terms of the TOEFL, probably the only objective measurement of student preparedness to enter full-time study currently available" (Saltzer, p.97).[1]

I refer those not familiar with the format of TOEFL to the Test and Score Manual (ETS (1981) pp. 7-8), but note also that the format is open to criticism on the grounds that extraneous variables and influences have not

---

1. Saltzer, however, mentions TOEFL's blind spots and dislikes the prospect of IELPs' curricula "being focused on the taking of the TOEFL rather than on language per se" (ibid.).

FL014587

been controlled sufficiently to differentiate the skills tested, so that the Structure section is to some extent a test of vocabulary, the Listening Comprehension section is to some extent a test of reading, etc. (criterion offered by Carrol, pp. 319-320).

Evidence for the growing popularity of TOEFL preparation can be found on the shelves of any campus or suburban bookstore, as well as by perusing the advertising material offered by IELPs to prospective students, either in their brochures and catalogues or in magazines such as "Study in the U.S.A." No doubt this fact alone is responsible for some of the heartsearching as to the propriety of offering TOEFL preparation in IELPS.

A more serious objection is that coaching for any purpose is a diversion of energy away from education. Coaching in its many forms has probably been with us -- and therefore suspect -- for centuries. Only recently, however, with the advent of IQ and other standardized tests supposedly measuring innate ability or aptitude developed over a lifetime has the opprobrium taken on the aspect of a crusade. A recent article (Bangert-Drowns et al) summarizes research on the effects of coaching for the SAT and some other standardized tests. With one-third of all high schools in the Northeast offering SAT preparation courses, and an estimated 50,000 students enrolling each year in commercial coaching schools, well-designed studies are beginning to emerge which show that "reasonably strong gains" can be expected on the SAT (ibid. p.81), and that gains of .5 standard deviations[2] have been attributed to coaching alone on other tests (ibid.). While the College Board as late as 1968 attacked this "obnoxious drill ... whereby the educational process [is being] unwillingly corrupted to gain ends ... not only unworthy but, ironically, unattainable" (quoted in U.S. Federal Trade Commission, p.32), the FTC report noted the credible results and called into question the fact that "research to date [focuses] on whether coaching is effective rather than why" (ibid. p.41).

It is noteworthy that ETS has retreated from its previous stance that coaching is ineffective. New ETS seminars and multi-media kits are "designed to help teachers and counselors prepare students for taking tests" (ETS (1983), p.20), and familiarize fourth-graders with standardized testing (ibid. p.4). In addition to the bulletins and handbooks traditionally provided to candidates for tests, there is now available "Testing American Style," informing foreign students of ways to ensure their best performance on standardized tests; a sample TOEFL; and the more elaborate "Understanding TOEFL: Test Kit 1." Along with the usual exhortations to study over a long period of time before attempting the test, this includes the carefully-worded statement that "Studying only the information in the kit will probably not improve your general ability to understand and use the English language. This can be achieved only by a regular program of language study" (ETS (1980), p.3). But can the student improve his score, regardless of his general ability? If not, why bother? And if so, then if "general ability"

---

2. Equivalent to 34 points on the TOEFL (ETS (1981), p.20).

does not improve, what does? What should the IELP's stance be towards coaching if even ETS admits that it can be efffective?

2. Program Attitudes

Some IELPs have adopted a favorable stance to TOEFL preparation. They may see it as a productive area, comparable to study skills, library use, or term paper writing. They may have noticed the "plateau effect" which can occur after a student has been in a program for several terms and begins to make little or no gain on TOEFL, the practice effect (given as 20 points over a 12-week instructional period: ETS (1981), p.25) not being so great after the first two or three times. They may combine either or both of these views with the "gut feeling" that the student deserves to be able to demonstrate what English language competency he has acquired to best advantage.

The hostile stance may be based on: (1) criticism (which may or may not be realistic or justified) of TOEFL, (2) the feeling that more can be done for the student by spending time on other things, (3) the view that there are no suitable or worthwhile materials available, or (4) the view so admirably summarized by the College Board report mentioned above. We will analyze below the view that TOEFL does not adequately measure English language proficiency. TOEFL is open to criticism in that it does not attempt to measure skills of production such as speaking and writing,[3] and also in that cultural factors which may impede the student's academic progress are not measured (Kaplan and Gue & Holdaway quoted in Homburg, p.31).

Some IELPs maintain a third stance which I have termed "acquiescent," possibly sympathizing with the views in the previous paragraph but still providing TOEFL preparation either as a service to their students or as an outright marketing gambit.

This spectrum of viewpoints reminds me of the stances adopted toward the provision of English for Specific Purposes (ESP) courses. An interesting cost-benefit analysis is provided by Strevens (p.119), counting as advantages of ESP (1) the focusing of time and effort on needed material only, (2) a higher rate of learning, motivation, and success, and (3) greater student (and therefore teacher) satisfaction. On the other hand, more preparation time and "superior, [and therefore] more 'expensive'" teachers are required. The similarities, in my view, are limited to the benefits, while the cost factor is not intrinsically present. IELPs appear to share this view, judging from the fact that TOEFL preparation courses, in addition to being more widely offered, are tending to become an integral part of the curriculum rather than either an optional extra course (often for an additional fee) or an elective.

3. Student Attitudes

---

3. But see Pike (1979) for a justification of the present Written Expression section compared with other possible measures.

Reference has already been made to the strong student demand for TOEFL preparation courses. I believe that this demand does not derive from any supposed desire on the part of the student to avoid meeting a test of his English language competency. If this were the case, the "hostile" position outlined above would be strengthened. Rather, this demand arises from a coalescence of personal, cultural, and career factors that will meet the definition of "need" as found in the phrase "needs analysis" (See e.g. Wilkins p.55 ff.).

It is acknowledged that students who may be deficient in English language skills still make academic progress once enrolled in regular academic courses: "deep conceptual ability" (Perkins & Pharis, p.146) may help to compensate in reading, and speaking ability, of course, along with writing skills beyond the sentence-unit, is not measured by TOEFL. When the student sees his fellow-nationals or colleagues making such progress, and he is still held back by the "TOEFL barrier," the rationale, purpose, and scope of the test is not clear to him (Angelis, p.104). "Inductive and logical reasoning abilities" (ibid. p.103) correlate strongly with success on reading comprehension tests such as that in TOEFL. An overt attempt to measure this might be unacceptable on grounds of "cultural bias," for which all TOEFL questions are screened (ETS (1981), p.8).

4. Uses of TOEFL

There is no shortage of literature on the alleged correlation between TOEFL scores and academic success. Chase (p.551) and Domino (Chase & Domino, p.552) summarized in 1972 the position which still holds sway today: "[TOEFL] may be of almost no value as a predictor of grade achievement; ... [this] may not be a reasonable criterion of the validity of TOEFL, since proficiency in English is a necessary but not sufficient prerequisite for academic achievement." In addition, as those who scored low on TOEFL were presumably refused entry to college, the criterion data from some of those who took the test are missing (Chai & Woehlke, pp. 16-17). The suggestion that "all potential candidates [be] admitted regardless of EFL test scores or other selection criteria ... [so that] some stronger relationship between test scores and academic performance [would] become apparent" (ibid. p.19) is not likely to be put to the test.

The most important use of TOEFL scores is as a measurement of English language proficiency. This is not as stable a characteristic as verbal or mathematical aptitude (ETS (1981), p.17). While Chase (p.550) found such "an obvious overlap in format between the Vocabulary and the English Structure subtests" that it was "not surprising to see that the intercorrelation between them [was] substantial," therefore minimizing the importance of part scores as opposed to overall TOEFL scores, a more recent study found support for "the position that a total TOEFL score may be less useful for assessment of language competence than the three section scores are, if interpreted in the light of the language background of the students" (Swinton & Powers, p.18). This, however, detracts from the usefulness of TOEFL as a worldwide standard of English proficiency. The same study makes a case for a separate

Vocabulary score (currently included in Section III with Reading Comprehension), as "vocabulary ... is acquired more rapidly than syntactic competence, and is thus more easily remediated ... if grammatical knowledge and reading comprehension are adequate" (ibid.).

In short, then, not only does "what ETS recommends be done with its test [vary] a great deal from what is actually being done with the TOEFL" (Homburg, p.30), but there is also doubt as to (1) which scores should be given greater stress (total or part), and (2) whether or not new part scores should be instituted. It should also be borne in mind that the potential impact on the decision-making process of the 14.6-point standard error of measurement[4] is often neglected. But the real question is to determine to what extent TOEFL is valid as a test of English language proficiency.

5. What Does TOEFL Measure?

To say that TOEFL measures English language proficiency is to imply (1) that a single, relevant standard exists against which the test-taker's proficiency can be measured; and (2) that language ability can be measured apart from communicative ability and social context. (These are two of Walters' "four basic assumptions of language testing", the others being "that language assessment instruments will accurately determine language ability [and] the effects of educational intervention" (Walters, pp. 294-297).)

The proficiency test is defined as a measurement of specific skills in the light of future performance of a particular task (Heaton p.164): much remains to be done in defining what the "tasks" involved in academic work are where foreign students are concerned. The layman's definition of "knowing" a language is invoked (Spolsky (1973), p.166) to mean performing in functional situations. Language use rather than language knowledge is being seen as more important: "social context selection rules are as necessary [as syntax] for linguistic competence" (Munby, p.9); "Language tests ... must take into account the full range of phenomena in communicative competence if language use is to be tested" (Briere p.329). It follows that we "have not yet found a way to characterize knowledge of a language with sufficient precision to guarantee[5] the validity of the items we include or the types of tests we use" (Spolsky (1972), p.348).

While discrete-point tests are appreciated for their ease of administration, objectivity, and reliability in scoring (Oller (1973), p.189), they tend to measure language knowledge rather than language use, inasfar as language

---

4. In other words, a 29.2-point spread on either side of the reported score for 95% of the examinees (ETS (1981), pp. 24-25).

5. ICP emphasis

6. Knowing when to "use" language in a given situation is a form of language

knowledge and language use can be separated from one another.[6] Discrete-point tests (wrongly) infer a degree of language proficiency from "testing a sample of surface features ... rather than tapping underlying linguistic competence" (Spolsky (1973), p.175). The current literature is full of references to the partitioning of language skill into "meaningful components which can be tested separately" (Oller & Hinofotis, p.14): "The notion of separate components of structure, vocabulary, and phonology finds no support [in the data given]" (ibid. p.23); "separate factors for [the four traditional] skill areas do not emerge" (Scholz et al, p.32). The alternative integrated-skills approach to testing has not yet progressed to the point at which it can challenge the primacy of the discrete-point test despite the undercutting of the theoretical defense of the latter, having "scarcely been defined except in opposition to discrete-point procedures" (Oller (1976), p.276), and, as with teaching methods, the solution would appear to be an "eclectic synthesis" (ibid. p.277) of the two. In the meantime the major rationale for accepting the TOEFL as a valid measure of English language proficiency rests on purely practical grounds.

The abandonment of the "Academic Lecture" subtest of TOEFL by 1979 was a clear indication of the preference of the test-makers for the discrete-point approach. While this subtest was criticized as measuring other factors besides listening skill (Chase, p.550) and as having too many technical handicaps to count as a "situation" (Cooper, p.71), it is noteworthy that no attempt was made to refine this one "situationalized" part of the test.

6. Responses

The long-term response to the inadequacy of the TOEFL will no doubt be the development of new instruments which will incorporate some elements of both approaches, testing for mastery of specifics and for communicative competence (See Valette, pp. 309-310). Upshur's widely-quoted observation that "trends in second language testing tend to follow trends in second language teaching" (p.177) leads him to believe that "as courses change to include communicative use of language so must language tests" (ibid.); even though this may be happening "on the ground," it may take much longer for ETS to adapt despite its previous willingness to evaluate alternative item formats within the discrete-point framework (See Pike, passim.).

In fairness to the interests of all of the groups identified by Clark (pp. 1-4) as involved in the testing process, an appropriate short-term response is to make sure that the student who takes TOEFL is as well prepared (in every sense of the word) as possible. This is a fourfold process.

(1): All literature dealing with standardized testing accepts that students should be familiarized with the format and procedures of the test. This may be limited to such "testwiseness" details as getting a good night's sleep, not looking around in the test room, and so forth, or enlarged to include the propriety of guessing under certain circumstances, scanning the questions

[6] "knowledge," just as pragmatic considerations may prove to be the only way to bridge the logical gap between linguistic "competence" and "performance" (See inter alia Derwing pp. 294-295).

before reading the passage, and marking questions to be returned to at the end in a certain order. This has been referred to as "test-taking strategy," in a limited sense of formal strategies rather than content-based strategies.

(2): Skill-building in regular classwork and in special "preparation" classes is necessary. A 30-point gain over a twelve-week period is attributed to "instructional effect" (ETS (1981), p.25). However, not all skills that are "built" actually reflect skills that are tested. A conscienscious preparation program will build onto the skills developed in the traditional "four skills" classwork, the refinements of those skills most closely identified with items which actually appear on the test. For example, many TOEFL preparation materials have Structure items such as: "By the time Juan gets home, his aunt _____ for Puerto Rico. [will leave\leaves\will have left\left]" In actual fact, TOEFL Structure items more closely correspond to the pattern: "_____ is indispensable to plant and animal life. [Nitrogen\It is nitrogen\That nitrogen\Although nitrogen]." Testing of manipulation of verb tenses is much less common than testing of clause structure in the first part of Section II (the Structure subtest).

(3): Practice is necessary, as a form of familiarization and as an anxiety reducer, but simply taking one practice test after another (even with explanatory answers), the "test-retest method," is not enough.[7] What can make a difference is to follow up skill-building and strategy (see below) sessions with practice items in test format, giving the student immediate feedback as to how well he has mastered the particular area. "Practice Tests" (part-tests of groups of items focusing on certain content areas) combined with full-length "composite" tests help restore the usefulness of the latter.[8]

(4): The most effective preparation method of all, especially when combined with the more widely practised methods above, is to train the students in test-taking techniques based on the design features of the test used by the test-makers. A list of the most common item types on each section and sub-section of the test, as established by analysis of actual TOEFLs, allows the construction of strategies designed to help the students apply their knowledge to the tasks found on the test. In some instances these strategies consist of an awareness of a frequently-used "trick" such as the use of distractors based on minimal-pair differences or sound similarities (Listening Comprehension section). Harris (p.57) criticizes attempts "to mislead examinees with distractors that look or sound like possible right answers," Spolsky et al (p.94) urge the avoidance of philological tricks, and

---

7. The "plateau effect" postulated above sees to it that the effect of constant practice diminishes with repetition.

8. Research still in progress with TOEFL preparation students at two Oregon programs.

BEST COPY AVAILABLE

Heaton notes that testers who "set deliberate traps or pitfalls for the unwary student" are themselves "caught out by ... trapping the more able student" (pp. 9-10). A familiarization with these strategies augments rather than detracts from (1) through (3) above. "Impressive results" were obtained from coaching materials developed by ETS researchers for specific item types included in the SAT item pool (Bangert-Drowns et al, p.81): this holds true for TOEFL item types also.

I am indebted to Pike (quoted in FTC, p.42) as to how coaching affects test scores: (1) the student is better prepared to demonstrate the underlying knowledge and reasoning abilities he has already acquired; (2) any gaps between that underlying base and the actual test content are filled in; (3) the student becomes familiar with the test format; (4) "test strategies" are explained (e.g. when to guess on the basis of limited information); (5) the anxiety level is reduced; and (6) the student will cope more efficiently with a timed situation.[9] The FTC report (ibid.) frets that the student may answer the questions correctly on the basis of "extraneous clues without any understanding of the skills being tested." They were appeased by ETS's reply that this would only rarely happen in a well-designed test, and stated that "[i]f coaching is serving a review function, then a well reviewed student's score is likely to be a better indicator of the student's actual potential" (ibid. p.44) with the caveat that the opposite would be true if the student learned skills but retained them only long enough to perform well on the test. It is up to each educator to decide this question for himself.

7. Conclusion

Limiting a teaching program to what was to be tested would be a strange set of priorities, to paraphrase Wilkins (p.82), but so would constructing a teaching program regardless of what was to be tested. My purpose has been to show that the most realistic and effective response to the situation students find themselves in with regard to the TOEFL is to provide the four test preparation methods outlined in Section 6 above in the appropriate proportions. The risks involved are that too much attention might be concentrated on questions of "testwiseness," and that the program might become too "test-centered," but also that what might be thought of as test preparation might simply be yet another structure review "dressed up" as TOEFL preparation.

The benefits, however, would be both in terms of answering a very deeply held concern on the part of the students and putting this concern in perspective. The IELP can therefore use the heightened motivation on the part of the students to enhance their English language competency and their

---

9. It is claimed that "adequacy of time is not a factor" in TOEFL because 80% [of a statistical group] completed all the questions in each section, and over 99% reached the three-quarter point (ETS (1981), p.24). However, we all know that students are encouraged to guess at the last moment and routinely leave no item unanswered.

career/academic opportunities.

If TOEFL does measure English language proficiency to the extent that it claims, increasing the students' knowledge of the type of discrete-point items tested increases their proficiency. If TOEFL does not measure English language proficiency to the extent that it claims, then the artificial barrier that it constitutes ought to be met frontally with all means available to the student and the IELP. I find the position totally illogical that "TOEFL is a bad test, therefore we should do nothing about it." The responsibility of the IELP is rather to determine how best to meet the student's need to become proficient in English and to display his proficiency to his best advantage. Surely the question is not whether this should be done, but rather how this can be done most effectively.

REFERENCES

Alderson, J.C. 1980. Native and Non-Native Speaker Performance on Cloze Tests. Occasional Papers on Linguistics, No. 6. Carbondale, Illinois: Southern Illinois University.

Angelis, Paul J. 1977. Language Testing and Intelligence Testing: Friends or Foes? Occasional Papers on Linguistics, No. 1. Carbondale, Illinois: Southern Illinois University.

Bangert-Drowns, Robert L., J.A. Kulik, and C.-L. Kulik. 1983. Synthesis of Research on the Effects of Coaching for Aptitude and Admissions Tests. Educational Leadership 1983, 41:4, 80-82.

Briere, Eugene. 1971. Are We Really Measuring Proficiency with Our Foreign-Language Tests? In: Teaching English as a Second Language: A Book of Readings, Allen, H.B. & Campbell, R.N., eds. 1972. New York: McGraw-Hill. Also in: Foreign Language Annals, 4: 385-391.

Carroll, John B. 1961. Fundamental Considerations in Testing for English Language Proficiency of Foreign Students. In: Teaching English as a Second Language: A Book of Readings, Allen, H.B. & Campbell, R.N., eds. 1972. New York: McGraw-Hill.

Chai, Sally H. & P.L. Woehlke. 1980. Predictive Ability of Standardized Tests of English as a Foreign Language. Occasional Papers on Linguistics, No. 6. Carbondale, Illinois: Southern Illinois University.

Chase, Clinton I. & G. Domino. 1972. A review in O.K. Buros (ed.), The Seventh Mental Measurements Yearbook. Highland Park, N.J.: The Gryphon Press.

Clark, John L.D. 1980. Foreign Language Testing: A Need for Communication. Occasional Papers on Linguistics, No. 6. Carbondale, Illinois: Southern Illinois University.

Cooper, Robert L. 1968. An Elaborated Language Testing Model. Language Learning, Special Issue 3 (Upshur, J.A. & Fata, J., eds.), pp. 57-65.

Derwing, Bruce L. 1973. Transformational Grammar as a Theory of Language Acquisition. Cambridge: Cambridge University Press.

Educational Testing Service, Information Services Division. Examiner, Special Issue, May 1983.

Educational Testing Service, TOEFL Program. 1981. TOEFL Test and Score Manual.

Educational Testing Service, TOEFL Program. 1980. Understanding TOEFL: Test Kit 1.

Gue, Leslie R. & Holdaway, E.A. 1973. English Proficiency Tests as Predictors of Success in Graduate Students in Education. Language Learning 1973, 23:1, 89-103.

Harris, David P. 1969. Testing English as a Second Language. New York: McGraw-Hill.

Heaton, J.B. 1975. Writing English Language Tests. London: Longman.

Homburg, Taco J. 1980. TOEFL and GPA: An Analysis of Correlations. Occasional Papers on Linguistics, No. 6. Carbondale, Illinois: Southern Illinois University.

Ingram, Elisabeth. 1968. Testing in the Context of a Language Learning Experiment. Language Learning, Special Issue 3 (Upshur, J.A. & Fata, J., eds.), pp. 147-157.

Munby, John. 1978. Communicative Syllabus Design. Cambridge: Cambridge University Press.

Oller, John W., Jr. 1973. Discrete-Point Tests Versus Tests of Integrative Skills. In: Focus on the Learner: Pragmatic Perspectives for the Language Teacher, Oller, J.W., Jr. & Richards, J.C., eds. Rowley, Mass.: Newbury House.

Oller, John W., Jr. 1976. Language Testing. In: A Survey of Applied Linguistics, Wardhaugh, R. & Brown, H.D., eds. Ann Arbor: University of Michigan Press.

Oller, John W., Jr. & Hinofotis, F.B. 1980. Two Mutually Exclusive Hypotheses About Second Language Ability: Indivisible or Partially Divisible Competence. In: Research in Language Testing, Oller, J.W. Jr. & Perkins, K., eds. Rowley, Mass.: Newbury House.

Perkins, Kyle & Pharis, K. 1980. TOEFL Scores in Relation to Standardized

Reading Tests. In: Research in Language Testing, Oller, J.W. Jr. & Perkins, K., eds. Rowley, Mass.: Newbury House.

Pike, Lewis W. 1979. An Evaluation of Alternative Item Formats for Testing English as a Foreign Language (TOEFL Research Report 2). Princeton, N.J.: Educational Testing Service.

Saltzer, Milton G. 1982. The Evaluation of an Intensive English Program. In: The Administration of Intensive English Language Progams, Barrett, R.P., ed. Washington, D.C.: National Association for Foreign Student Affairs.

Scholz, George, et al. 1980. Is Language Ability Divisible or Unitary? In: Research in Language Testing, Oller, J.W. Jr. & Perkins, K., eds. Rowley, Mass.: Newbury House.

Spolsky, Bernard. 1972. Language Testing—The Problem of Validation. In: Readings on English as a Second Language, Croft, K., ed. Cambridge, Mass.: Winthrop. Also in TESOL Quarterly 2: 88-94.

Spolsky, Bernard. 1973. What Does It Mean to Know a Language, or How Do You Get Someone to Perform His Competence? In: Focus on the Learner: Pragmatic Perspectives for the Language Teacher, Oller, J.W., Jr. & Richards, J.C., eds. Rowley, Mass.: Newbury House.

Spolsky, Bernard, et al. 1968. Preliminary Studies in the Development of Techniques for Testing Overall Second Language Proficiency. Language Learning, Special Issue 3 (Upshur, J.A. & Fata, J., eds.), pp. 79-98.

Strevens, Peter. 1980. Teaching English as an International Language. Oxford: Pergamon Press.

Swinton, Spencer S. & Powers, D.E. 1980. Factor Analysis of the Test of English as a Foreign Language for Several Language Groups (TOEFL Research Report 6). Princeton, N.J.: Educational Testing Service.

United States Federal Trade Commission. 1981. Effects of Coaching on Standardized Admission Examinations. (NTIS PB-296 196) ERIC ED 204-391.

Upshur, John A. 1969. Productive Communication Testing: Progress Report. In: Focus on the Learner: Pragmatic Perspectives for the Language Teacher, Oller, J.W., Jr. & Richards, J.C., eds. 1973. Rowley, Mass.: Newbury House.

Valette, R.M. 1977. Modern Language Testing: A Handbook. 2nd ed. New York: Harcourt, Brace.

Walters, Joel. 1980. Language Variation in the Assessment of Communicative Competence. Occasional Papers on Linguistics, No. 6. Carbondale, Illinois: Southern Illinois University.

Wilkins, D.A. 1976. Notional Syllabuses. Oxford: Oxford University Press.

Hale, Gordon A., C.W. Stansfield, and R.P. Duran (1984). Summaries of Studies Involving the Test of English as a Foreign Language, 1963-1982 (TOEFL Research Report 16). Princeton, N.J.: Educational Testing Service.

Hale, Gordon A., C.W. Stansfield, and R.P. Duran (1984). A Comprehensive TOEFL Bibliography, 1963-1982. Modern Language Journal 1984, 67-4.